06 | 帧内预测：如何减少空间冗余？

2021-12-03 李江

课程介绍 >

讲述：李江
时长 15:01　大小 13.76M

你好，我是李江。

前面几节课我们一起讨论了视频编码的原理以及编码码流结构，从今天开始我们一起来聊一聊具体的编码模式。

我们知道在视频编码时主要需要减少 4 个冗余，包括：空间冗余、时间冗余、视觉冗余和信息熵冗余。其中空间冗余就是通过帧内预测的方式来实现的。那帧内预测到底是怎么做到减少空间冗余的呢？

其实一般来说，**一幅图像中相邻像素的亮度和色度信息是比较接近的，并且亮度和色度信息也是逐渐变化的，不太会出现突变**。也就是说，**图像具有空间相关性**。帧内预测就是利

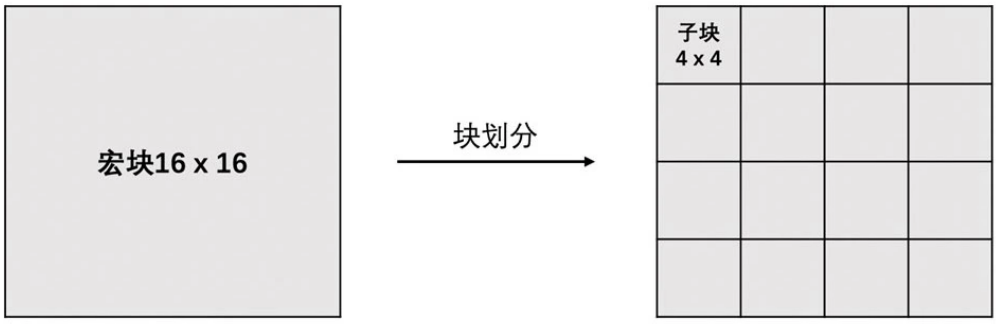
用这个特点来进行的。即**帧内预测通过利用已经编码的相邻像素的值来预测待编码的像素值，最后达到减少空间冗余的目的**。

这里需要注意的一个点就是，我们是通过已经编码了的像素值去预测待编码的像素值。你可能会问，已经编码了的像素值变成码流了，不再是一个个像素了，怎么去预测待编码的像素呢？其实已经编码了的像素是会重建成重建像素，用来做之后待编码块的参考像素的。你可以认为是**已经编码的块会解码成像素用来做参考像素**。

好了，我们已经知道了帧内预测是怎么减少空间冗余的，那我们还有一个最重要的问题要解答。既然帧内预测是通过相邻像素来预测待编码像素的，那它到底是怎么预测的呢？下面我们就以 H264 标准来讲解每种帧内预测模式是怎么做预测的。

不同块大小的帧内预测模式

我们在视频编码原理的那节课里面讲过，视频编码是以块为单位进行的。**在 H264 标准里面，块分为宏块和子块。宏块的大小是 16 x 16（YUV 4:2:0 图像亮度块为 16 x 16，色度块为 8 x 8）**。在帧内预测中，亮度宏块可以继续划分成 16 个 4 x 4 的子块。因为图像中有的地方细节很多，我们需要划分成更小的块来做预测会更精细，所以会将宏块再划分成 4 x 4 的子块。如下图所示：



极客时间

帧内预测是根据块的大小分为不同的预测模式的。还有一个点就是亮度块和色度块的预测是分开进行的。这里我给你总结一下，主要有以下 3 点规则：

1. **宏块大小是 16 x 16，其中亮度块为 16 x 16，色度块为 8 x 8**；
2. **帧内预测中亮度块和色度块是分开独立进行预测的**，即亮度块参考已编码亮度块的像素，而色度块参考已编码色度块的像素；

3. **16 x 16 的亮度块可以继续划分成 16 个 4 x 4 的子块。**

所以，我们在实际帧内预测的时候就会分为：4 x 4 亮度块的预测、16 x 16 亮度块的预测、8 x 8 色度块的预测（注意亮度 8 x 8 模式和 I_PCM 模式很少使用，我们这里不做讨论）。下面我们就开始一个个展开来讨论吧。

4 x 4 亮度块的帧内预测模式

因为 4 x 4 的块帧内预测模式最多，并且基本包含亮度 16 x 16 和色度 8 x 8 的模式，所以我们接下来先从 4 x 4 亮度块的帧内预测模式讲起。

4 x 4 亮度块的帧内预测模式总共有 9 个。其中有 **8 种方向模式和一种 DC 模式**，且方向模式指的是预测是有方向角度的。好了，那我们接下来就聊聊这 9 种模式吧。

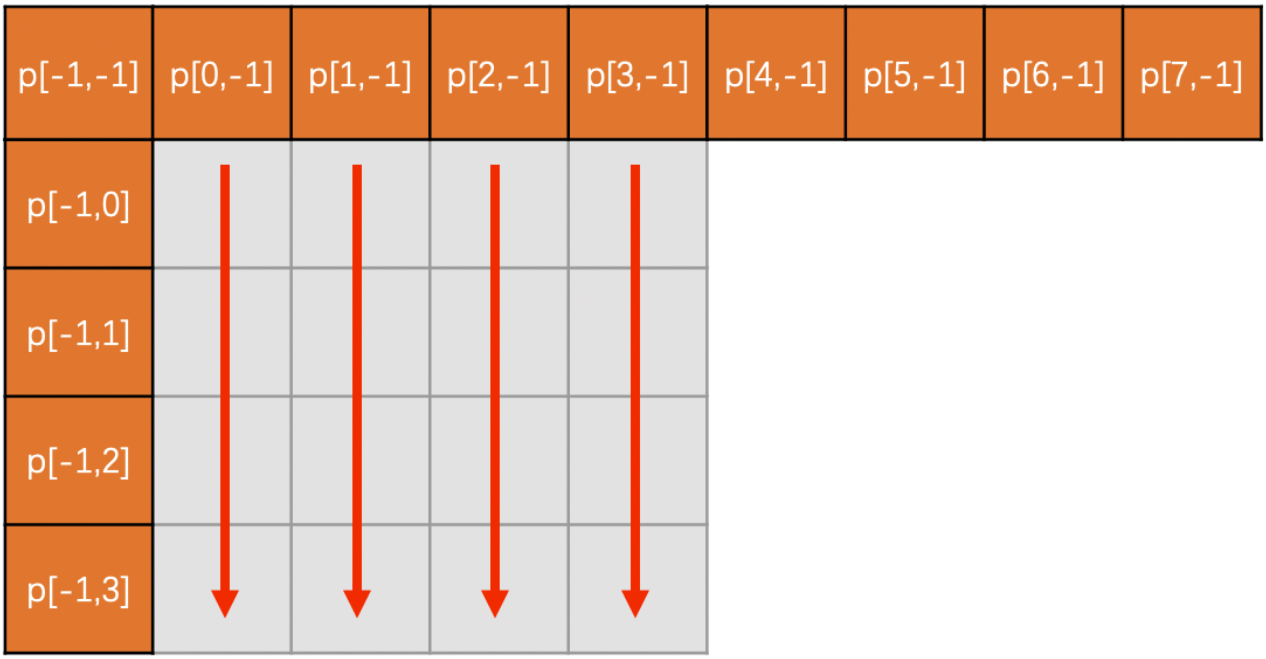
1. Vertical 模式

Vertical 模式就是指，当前编码亮度块的每一列的像素值，都是复制上边已经编码块的最下面那一行的对应位置的像素值。这句有点拗口，别急，等会儿我会给出计算方法和示意图，让你能很容易地理解它。

Vertical 模式得到的预测块同一列中的像素值都是一样的。该模式得到的块就叫做 Vertical 预测块。注意，**该模式只有在上边块存在的时候才可用，如果不存在则该模式不可用。**比如图像最上边的块就没有可参考的块存在。

该模式下预测块像素计算方法如下：

示意图如下：

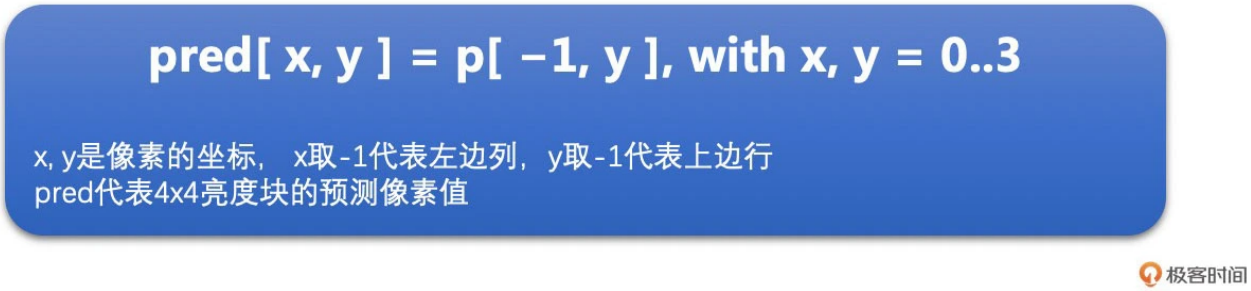


编码块（4x4）
灰色部分：编码块
橙色部分：左边、上边、左上角对角的参考像素

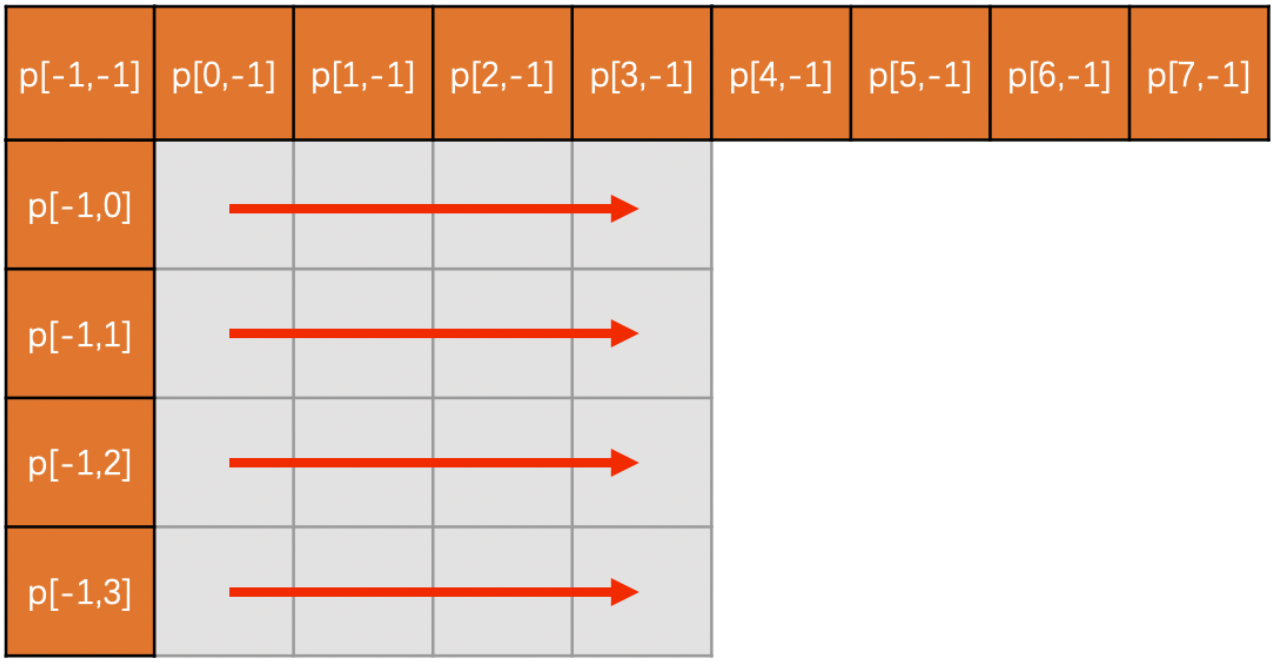
2. Horizontal 模式

Horizontal 模式就是指，当前编码亮度块的每一行的像素值，都是复制左边已经编码块的最右边那一列的对应位置的像素值（同样地，下面有计算方法和示意图）。**Horizontal 模式得到的预测块同一行的像素值都是一样的**，该模式得到的块就叫做 Horizontal 预测块。注意，**该模式只有在左边块存在的时候才可用，如果不存在则该模式不可用**。比如图像最左边的块就没有可参考的块存在。

该模式下预测块像素计算方法如下：



示意图如下：



编码块（4x4）
灰色部分：编码块
橙色部分：左边、上边、左上角对角的参考像素

3. DC 模式

DC 模式就是指，当前编码亮度块的每一个像素值，是上边已经编码块的最下面那一行和左边已编码块右边最后一列的所有像素值的平均值。注意，**DC 模式预测得到的块中每一个像素值都是一样的。**DC 模式得到的块就叫做 DC 预测块。

根据上边块和左边块是不是存在，该模式下预测块像素的计算方法分为以下四种情况：

第一种情况就是上边块和左边块都存在的话，预测块像素计算如下：

$$pred[x, y] = (p[0, -1] + p[1, -1] + p[2, -1] + p[3, -1] + p[-1, 0] + p[-1, 1] + p[-1, 2] + p[-1, 3] + 4) >> 3$$

第二种情况是上边块不存在，左边块存在的话，预测块像素计算如下：

$$pred[x, y] = (p[-1, 0] + p[-1, 1] + p[-1, 2] + p[-1, 3] + 2) >> 2$$

第三种情况是左边块不存在，上边块存在的话，预测块像素计算如下：

$$pred[x, y] = (p[0, -1] + p[1, -1] + p[2, -1] + p[3, -1] + 2) >> 2$$

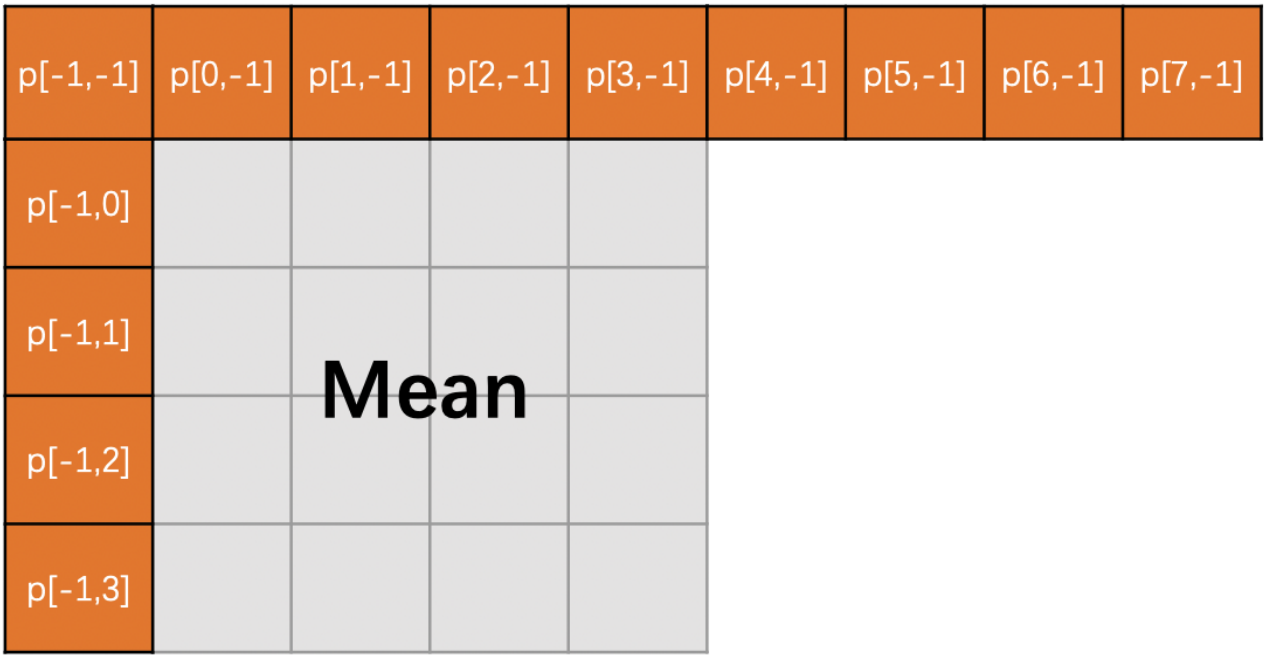
第四种情况是上边块和左边块都不存在，预测块像素计算如下：

$$pred[x, y] = (1 << (bit\_depth - 1))$$

// bit_depth是位深，我们用到的都是8位

极客时间

示意图如下：

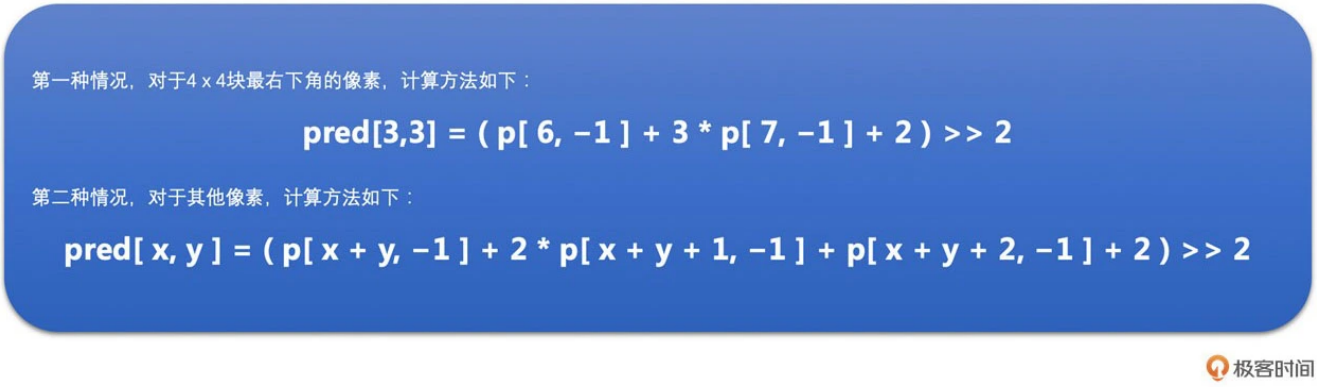


编码块（4x4）
灰色部分：编码块
橙色部分：左边、上边、左上角对角的参考像素

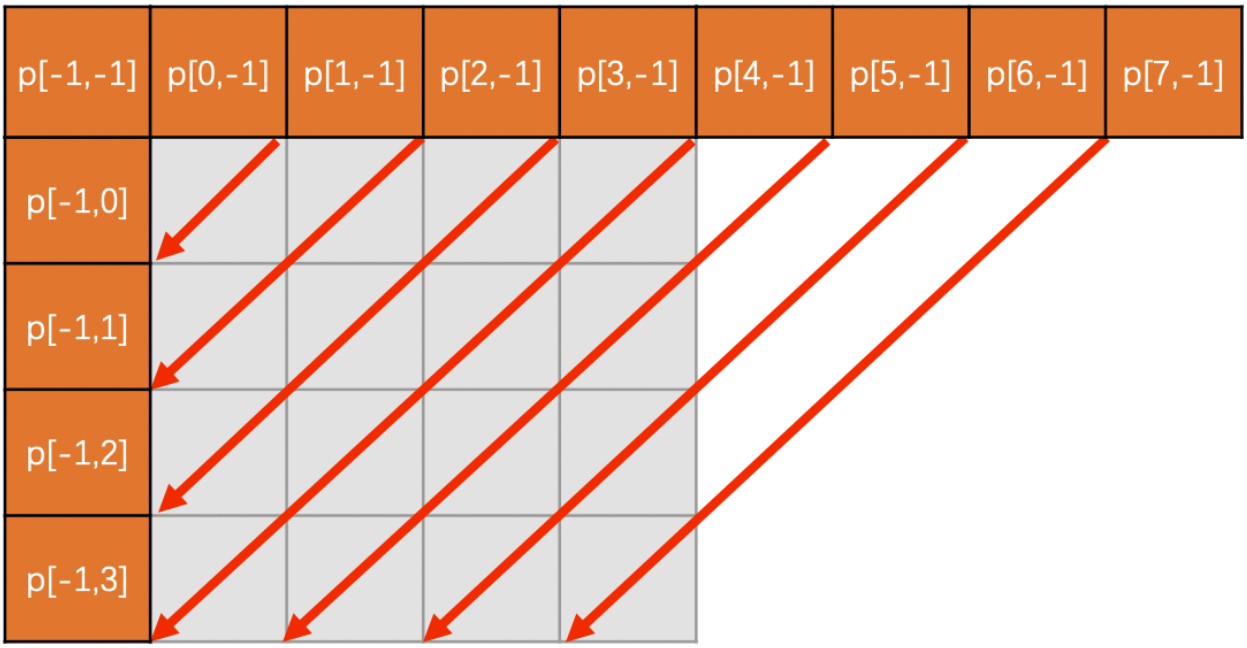
4. Diagonal Down-Left 模式

Diagonal Down-Left 模式是上边块和右上块（上边块和右上块有可能是一个块，因为可能是一个 16 x 16 的亮度块，意思理解就可以）**的像素通过插值得到。如果上边块和右上块不存在则该模式无效。**

该模式下预测块计算的方法分为下面两种情况：



示意图如下：



编码块（4x4）
灰色部分：编码块
橙色部分：左边、上边、左上角对角的参考像素

5. Diagonal Down-Right 模式

Diagonal Down-Right 模式需要通过上边块、左边块和左上角对角的像素通过插值得到。如果这三个有一个不存在则该模式无效。

该模式下预测块计算的方法有以下三种情况：

第一种情况，对于x > y，计算方法如下：

$$pred[x, y] = (p[x - y - 2, -1] + 2 * p[x - y - 1, -1] + p[x - y, -1] + 2) >> 2$$

第二种情况，对于x < y，计算方法如下：

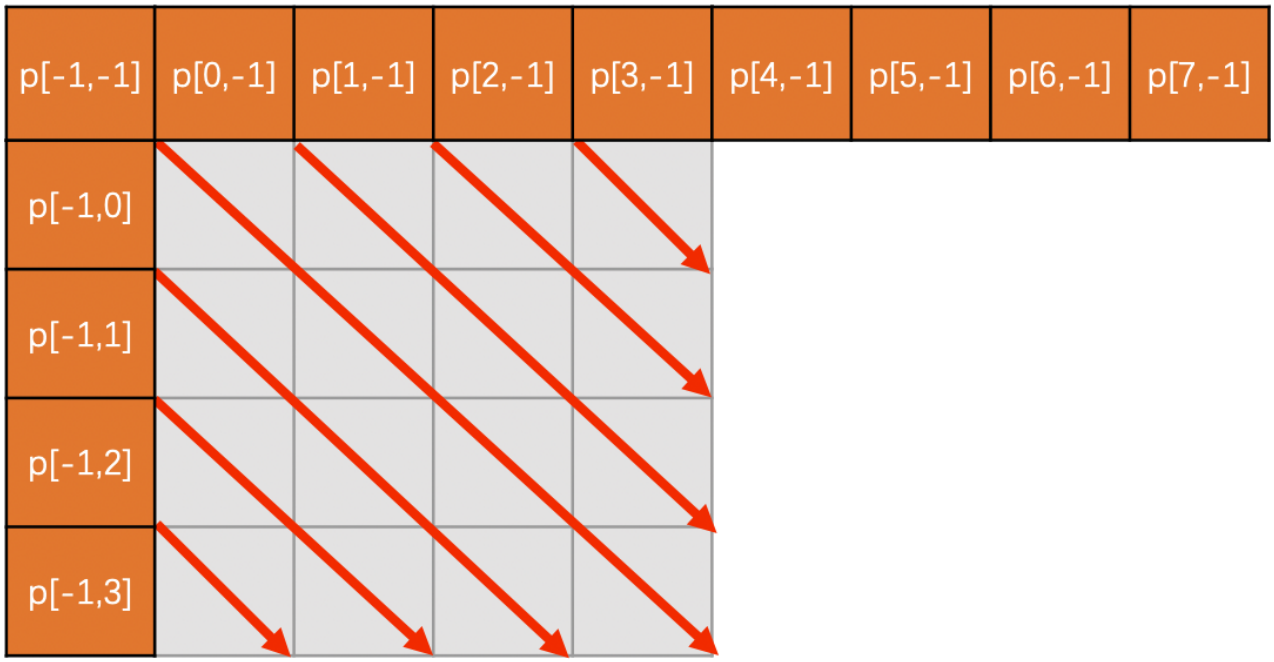
$$pred[x, y] = (p[-1, y - x - 2] + 2 * p[-1, y - x - 1] + p[-1, y - x] + 2) >> 2$$

第三种情况，对于x = y，计算方法如下：

$$pred[x, y] = (p[0, -1] + 2 * p[-1, -1] + p[-1, 0] + 2) >> 2$$

极客时间

示意图如下：

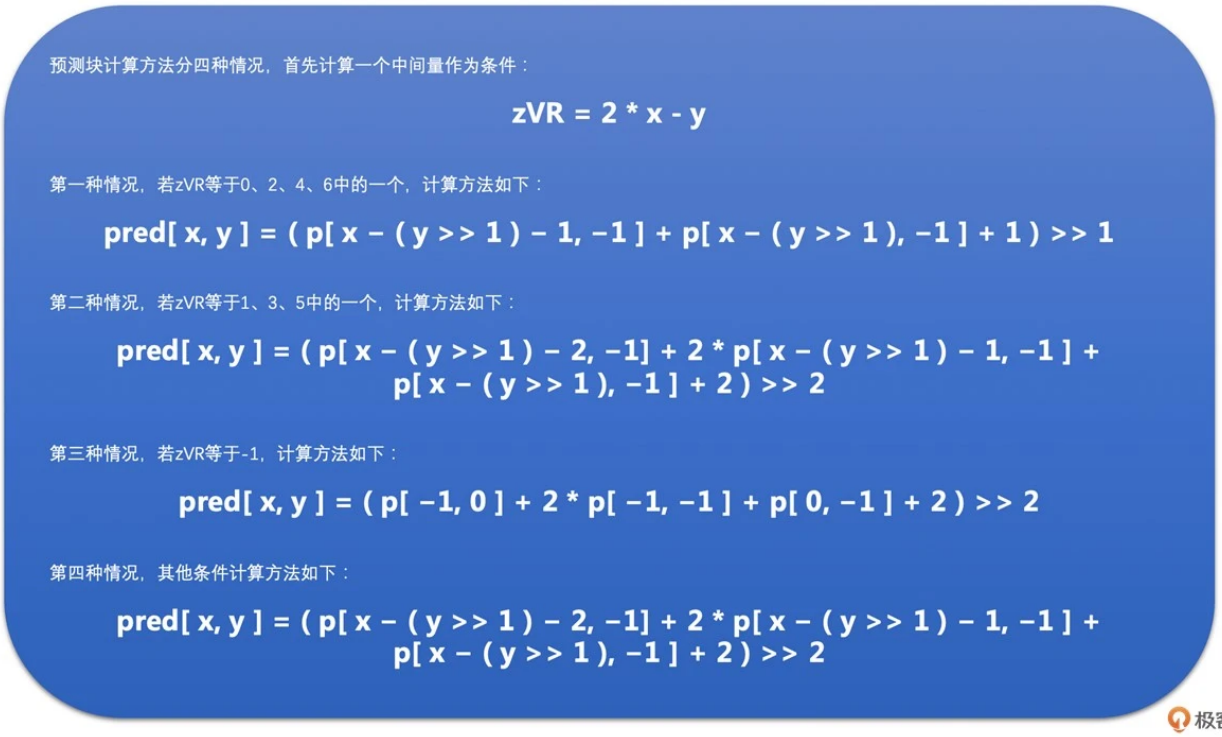


编码块（4x4）
灰色部分：编码块
橙色部分：左边、上边、左上角对角的参考像素

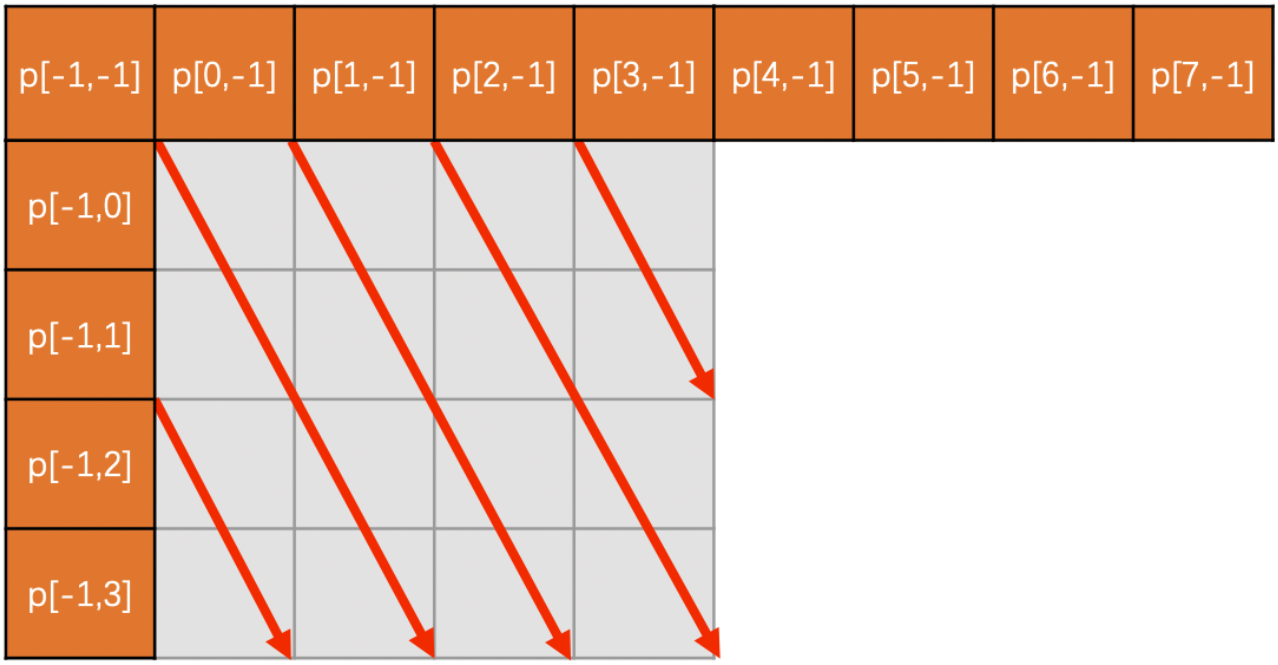
6. Vertical-Right 模式

Vertical-Right 模式是需要通过上边块、左边块以及左上角对角的像素插值得到的。必须要这三个都有效才能使用，否则该模式无效。

该模式下预测块计算的方法有以下四种情况：


预测块计算方法分四种情况，首先计算一个中间量作为条件：
zVR = 2 * x - y
第一种情况，若zVR等于0、2、4、6中的一个，计算方法如下：
pred[x, y] = (p[x − (y >> 1) − 1, −1] + p[x − (y >> 1), −1] + 1) >> 1
第二种情况，若zVR等于1、3、5中的一个，计算方法如下：
pred[x, y] = (p[x − (y >> 1) − 2, −1] + 2 * p[x − (y >> 1) − 1, −1] +
p[x − (y >> 1), −1] + 2) >> 2
第三种情况，若zVR等于-1，计算方法如下：
pred[x, y] = (p[−1, 0] + 2 * p[−1, −1] + p[0, −1] + 2) >> 2
第四种情况，其他条件计算方法如下：
pred[x, y] = (p[x − (y >> 1) − 2, −1] + 2 * p[x − (y >> 1) − 1, −1] +
p[x − (y >> 1), −1] + 2) >> 2
极客时间

示意图如下：

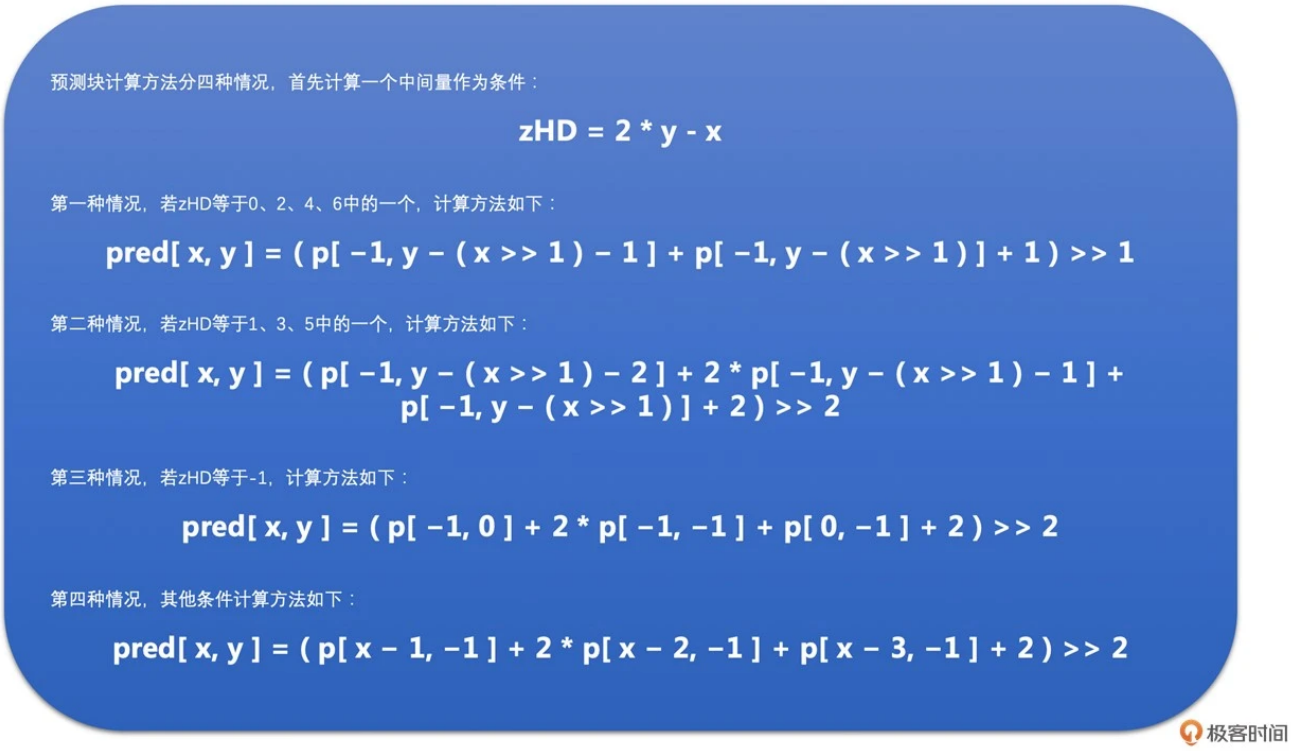


编码块（4x4）
灰色部分：编码块
橙色部分：左边、上边、左上角对角的参考像素

7. Horizontal-Down 模式

Horizontal-Down 模式需要通过上边块、左边块以及左上角对角的像素插值得到。必须要这三个都有效才能使用，否则该模式无效。

该模式下预测块计算的方法有以下四种情况：

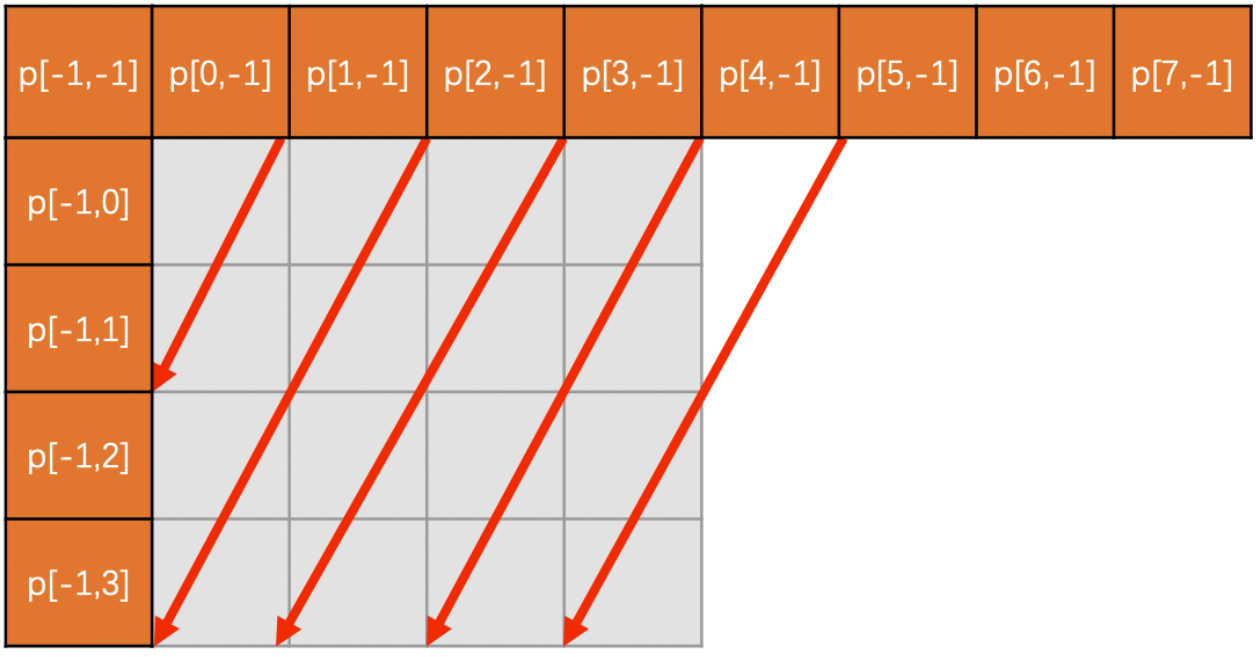


示意图如下：

8. Vertical-Left 模式

Vertical-Left 模式是需要通过上边块和右上块（上边块和右上块有可能是一个块，因为可能是一个 16 x 16 的亮度块，意思理解就可以）**最下面一行的像素通过插值得到**。如果这两种块不存在则该模式不可用。该模式下预测块的计算方法有下面两种情况：

示意图如下：



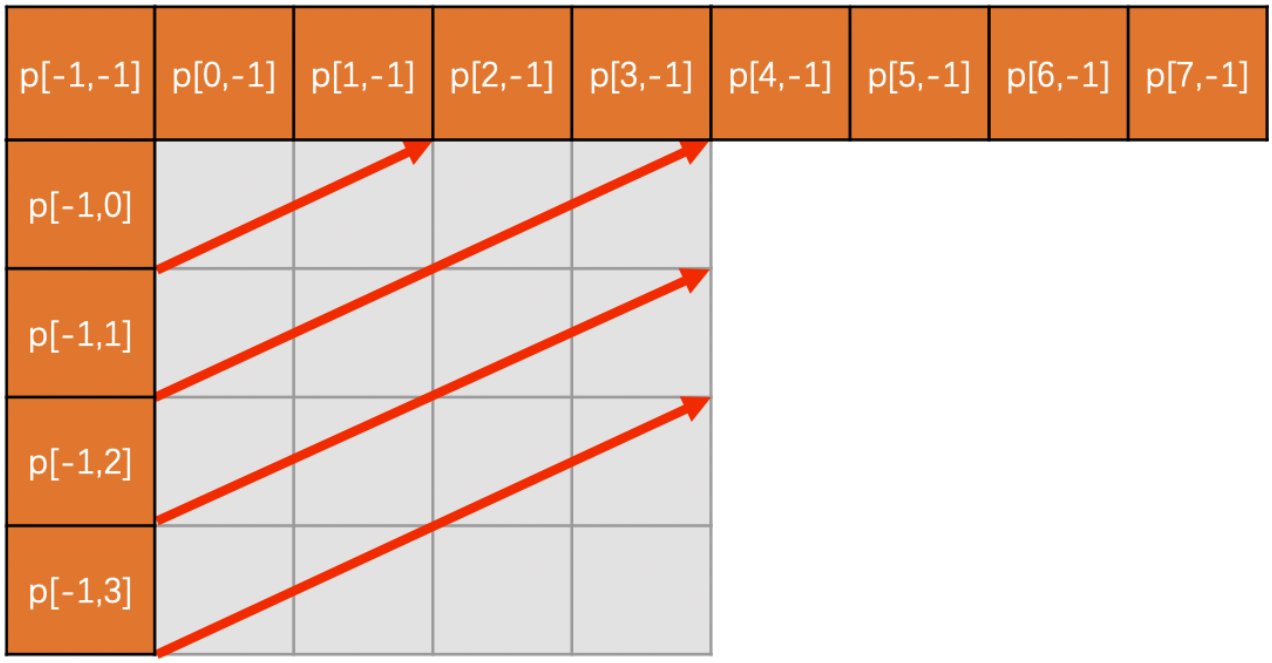
编码块（4x4）
灰色部分：编码块
橙色部分：左边、上边、左上角对角的参考像素

9. Horizontal-Up 模式

Horizontal-Up 模式是需要通过左边块的像素通过插值得到的。如果左边块不存在，则该模式不可用。

该模式下预测块的计算方法有下面四种情况：

示意图如下：

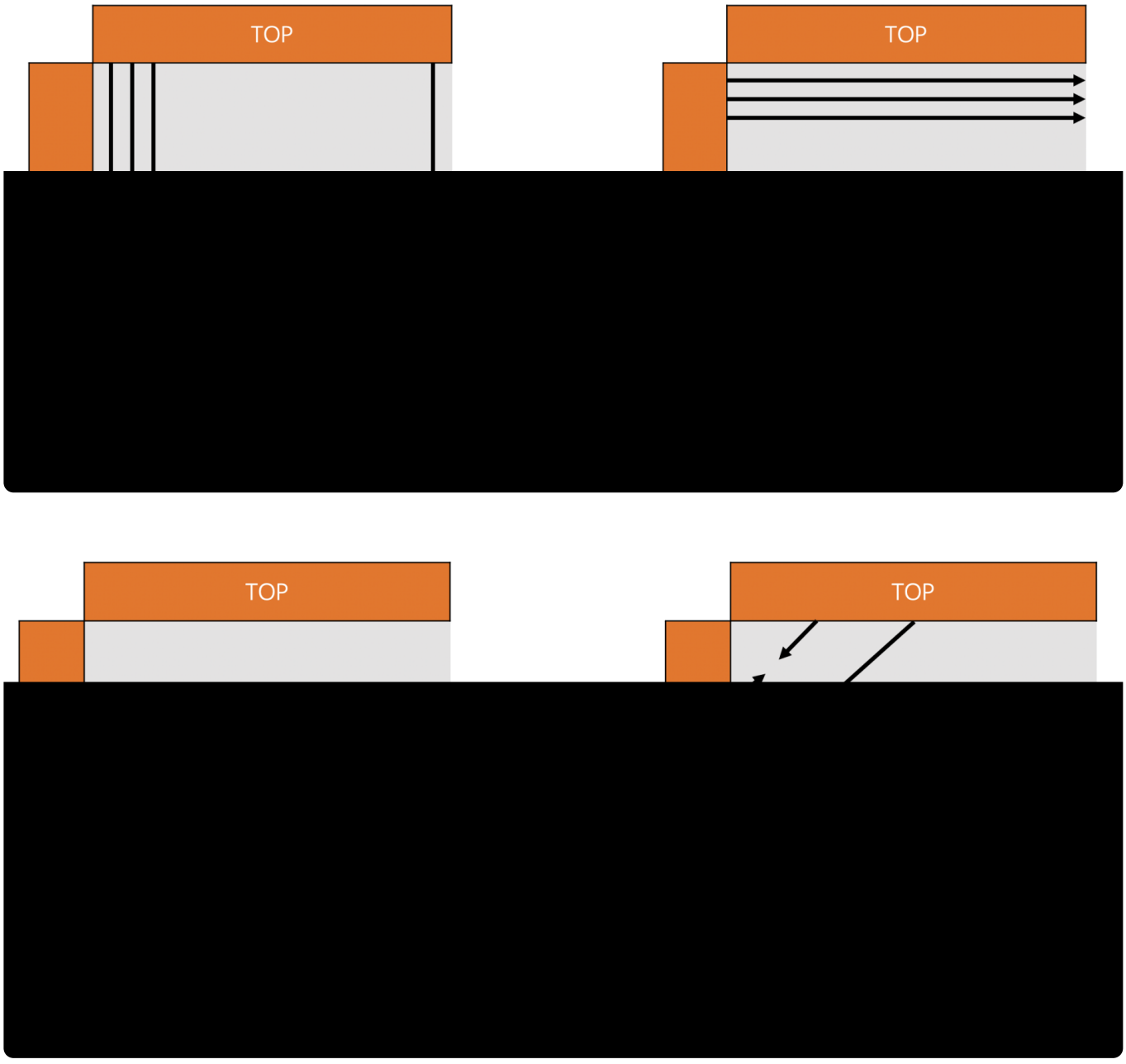


编码块（4x4）
灰色部分：编码块
橙色部分：左边、上边、左上角对角的参考像素

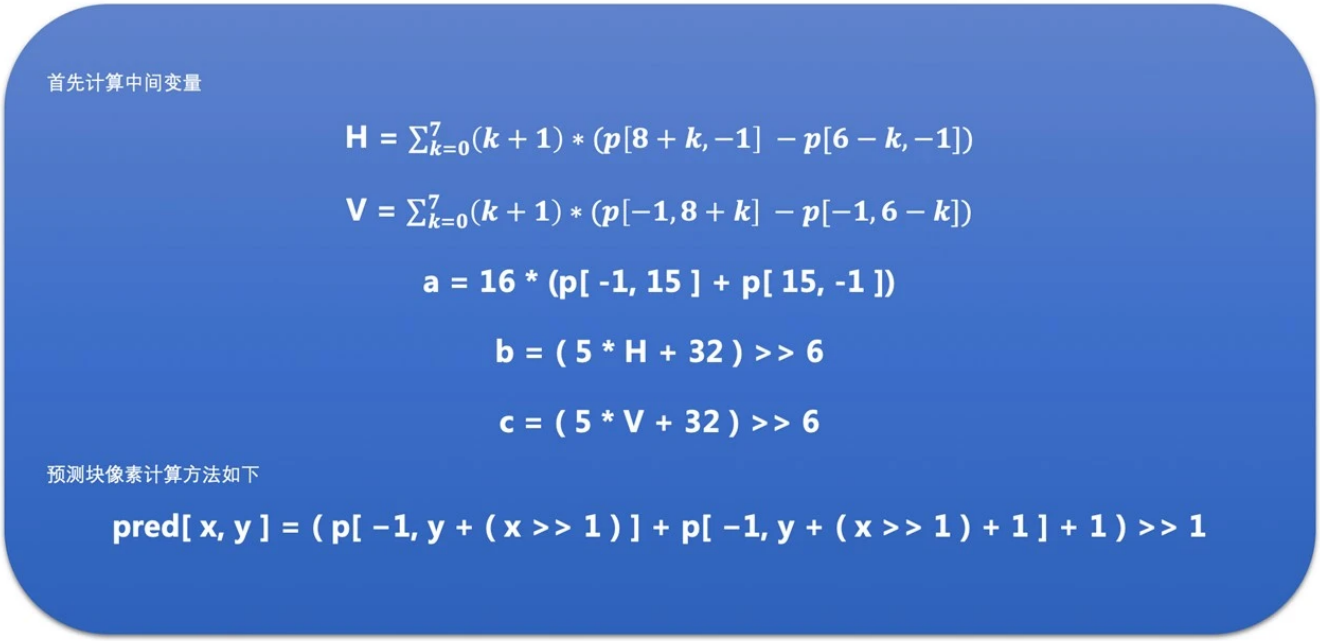
16 x 16 亮度块的帧内预测模式

上面我们讲了 4 x 4 亮度块的 9 种预测模式，接下来我们来聊一下 16 x 16 亮度块的预测模式。

16 x 16 亮度块总共有4种预测模式。它们分别是Vertical模式，Horizontal模式、DC 模式和Plane 模式，前面三种模式跟 4 x 4 的原理是一样的这里就不重复讲解了。16 x 16 亮度块的帧内预测的 4 种模式的示意图如下：



接下来我们介绍一下 Plane 模式。这种模式相比前面三种模式稍微复杂一些，但是基本原理都差不多。**Plane 预测块的每一个像素值，都是将上边已编码块的最下面那一行，和左边已编码块右边最后一列的像素值经过下面公式计算得到的。**



8 x 8 色度块的帧内预测模式

上面我们讲述了亮度块的帧内预测模式，接下来我们简单介绍一下色度块的帧内预测模式。

8 x 8 色度块的帧内预测模式跟 16 x 16 亮度块的是一样的，也是总共有 4 种，分别为 DC 模式、Vertical 模式，Horizontal 模式、Plane 模式。**与 16 x 16 亮度块不同的是，块大小不同，所以参考像素值数量会不同。**但是基本是一致的，这里就不重复了。

帧内预测模式的选择

现在我们学习了这么多的模式，而每一个块却只能有一种帧内预测模式。那我们怎么确定一个块到底使用哪种模式呢？我们这边先把思路讲一讲，具体细节不展开。等到你把编码基础弄明白了之后可以阅读一下 x264 的代码，里面有关于具体如何去选择模式的方法。

以亮度块为例。一个 16 x 16 的亮度块，我们可以不划分，直接使用 4 种 16 x 16 的帧内预测模式，最多得到 4 种预测块；也可以划分成 16 个 4 x 4 的子块，每一个子块最多有 9 种帧内预测模式。

对于每一个块或者子块，我们可以得到预测块，再用实际待编码的块减去预测块就可以得到残差块。主要有下面 3 种方案来得到最优预测模式：

第一种方案，先对每一种预测模式的残差块的像素值求绝对值再求和，称之为 cost，然后取其中残差块绝对值之和也就是 cost 最小的预测模式为最优预测模式。

第二种方案，对残差块先进行 Hadamard 变换（在 DCT 变换和量化那节课中会介绍），变换到频域之后再求绝对值求和，同样称为 cost，然后取 cost 最小的预测模式为最优预测模式。

第三种方案，也可以对残差块直接进行 DCT 变换量化熵编码，计算得到失真大小和编码后的码流大小，然后通过率失真优化（作为课外内容自行学习，这里不展开讨论）的方法来选择最优预测模式。
为了让你更容易理解第三种方案，这里我稍微简单介绍一下率失真优化的思想。

我们知道预测之后经过 DCT 变换再量化会丢失高频信息。一般来说 QP 越大，丢失的信息越多，失真就越大，但是码流大小也越小；反之，QP 越小，丢失的信息越少，但是码流大小就越大。这是一个跷跷板。**我们一般会在失真和码流大小之间平衡，尽量找到在一定码率下，失真最小的模式作为最优的预测模式，这就是率失真优化的思想。**

其实还有很多不同的方案，比如有的为了加速模式选择的过程，率失真计算的时候，只会进行 DCT 变换和量化，不会进行熵编码。码流大小直接通过 QP 值估算或者使用预测模式的大小来代替。这些方案都可以，具体看编码器的实现。一般来说，选择过程越精细效果越好，但是速度会越慢。

通过上面讲的这些方法我们找到了每一个 4 x 4 块的最优模式之后，将这 16 个 4 x 4 块的 cost 加起来，与 16 x 16 块的最小 cost 对比，选择 cost 最小的块划分方式和预测模式作为帧内预测模式。

小结

好了，今天的课到这里就要结束了。我们来回顾一下这节课的主要内容。

我们主要是一起聊了聊帧内预测。我们先详细讲述了有哪几种预测模式，以及各种预测模式的预测块的计算方法。同时强调一点，就是有的模式在参考像素不存在的时候是无效的，也就是不能使用的。这一点需要注意，并不是列举出来的每一个模式都是可以使用的。

为了方便你记忆，下面我把每一种块和子块支持的预测模式用一个表格来总结一下。

块类型	帧内预测模式
亮度4 x 4	Vertical模式
	Horizontal模式
	DC模式
	Diagonal Down-Left模式
	Diagonal Down-Right模式
	Vertical-Right模式
	Horizontal-Down模式
	Vertical-Left模式
	Horizontal-Up模式
亮度16 x 16	Vertical模式
	Horizontal模式
	DC模式
	Plane模式
色度8 x 8	DC模式
	Vertical模式
	Horizontal模式
	Plane模式

最后，我们还简单介绍了一下预测模式的选择方法，主要有计算残差块绝对值之和、将残差块做 Hadamard 变换之后再求和、率失真优化等几种方案来得到 cost，然后我们取 cost 最小的模式作为帧内预测模式。

思考题

在 H264 标准里面，视频的第一帧的第一个块应该怎么选择预测模式呢？

你可以把你的答案留在评论区，和我一起探讨、交流，看看自己的理解是不是正确。下节课再见。

分享给好友，一起充电升级

赞 0 提建议

© 版权归极客邦科技所有，未经许可不得传播售卖。 页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 05｜码流结构：原来你是这样的H264

精选留言

写留言

> 由作者筛选后的优质留言将会公开显示，欢迎踊跃留言。